Probability in the Engineering and Informational Sciences, **33**, 2019, 487–499.
doi:10.1017/S0269964818000505

A SPECTRUM OF SERIES–PARALLEL GRAPHS WITH MULTIPLE EDGE EVOLUTION

HOSAM M. MAHMOUD
Department of Statistics,
The George Washington University, Washington, D.C., 20052, USA
E-mail: hosam@gwu.edu

We discuss a rich family of directed series–parallel (SP) graphs grown by the simultaneous random series or parallel development of multiple edges. The family portrays a spectrum that spans a wide range of SP graphs: from simple models, where only as few as one edge is chosen for evolution at each discrete point in time, to complex hierarchical lattice networks grown by a take-all strategy, where all the edges in the existing network are developed.

The family of SP graphs we discuss is grown from an initial seed graph with τ_0 edges under an arbitrary *building sequence*, $\{k_n\}_{n=1}^{\infty}$, of nonnegative integers (with $k_n \le \tau_0 + \sum_{i=1}^{n} k_i$, for arbitrary $\tau_0 \ge 1$), that specifies the number of edges subjected to evolution at time n. We study the average north polar degree and show that we can go beyond averages to strong laws. We also find the exact average number of critical edges. The asymptotics of the critical edges are facilitated under the regularity condition that $k_n / \sum_{i=1}^{n} k_i$ converges to a constant (as $n \to \infty$), a natural condition easily met by practical strategies, such as single-edge evolution and take-all choice, and much in between.

Keywords: flow in graph, martingale, network, random graph, recurrence

Mathematics Subject Classifications: Primary: 60C05, 05C80, 90B15;
Secondary: 60F15, 05C21.

1. INTRODUCTION

Random graphs are versatile models that capture many real-world applications in algorithmics and communication networks, among countless other applications. For the two named applications, the reader can see, for example [3,8,12,13,17], for a myriad of random tree applications in computer science, and [1,18] for numerous applications in communication networks, among many other types of hardwired and virtual networks.

An algorithmically important family of structures is the series–parallel (SP) graphs that represent flow in networks (of commercial goods, e.g.) from a source (a factory, e.g.) to a sink (a market, e.g.).

Not much work has been done in the area of probabilistic analysis of SP graphs. In fact, there are not many known probability models on these graphs. This motivates us to build more probability models for SP graphs.

© Cambridge University Press 2019 0269-9648/18 $25.00

https://doi.org/10.1017/S0269964818000505 Published online by Cambridge University Press

Among the few known probability models of randomness on SP graphs are the uniform model, where all SP networks of a certain size are equally likely [2,4], the hierarchical lattice model [7], the incremental models for unrestricted and binary SP graphs (with node outdegrees restricted to 2) [14,15]. The source [10] investigates the behavior of binary SP graphs under a uniform probability distribution.

In this manuscript, we consider one stochastic growth model, which may cover a wide variety of realistic applications. In this paragraph, we highlight the model and mention possible merits, and in Section 2, we specify it formally. The model is based on choosing a number of edges $k_n \geq 0$ from an SP graph available at time $n-1$, and letting the selected edges evolve independently via serialization and parallelization operations. Instances of this family of graphs have been looked at: In [14] a slow incremental model selects one edge at each step of growth and subjects it to a probabilistic SP evolution ($k_n = 1$). At the other end of the spectrum, the hierarchical lattice model [7] lets every edge in the graph experience its own independent evolution starting at the complete graph on one edge ($k_n = 2^{n-1}$). The slow model [14] may suit market flow applications and road networks, whereas the hierarchical lattice model [7] represents very fast growing networks, such as some of the social networks evolving today. The model we suggest here is meant to fill the spectrum with various growth speeds, and may thus capture a wider variety of modern needs.

Note that the allowable choice $k_n = 0$ corresponds to no edge selection and subsequently no evolution at time n. This device can be used to slow down the evolution, which may suit some applications.

2. MULTI-EDGE GROWTH OF SP GRAPHS

We start with some formal definitions. As is customary, we use the notation K_n for the complete graphs on n vertices.

One popular definition of SP graphs comes from an inductive algorithmic construction.[1] In this variant, the smallest SP graph is K_2. The two vertices are called *Poles*. We depict K_2 by a drawing that puts it in a vertical position on a page. The top pole is called the *North Pole* (N), and the bottom one is called the *South Pole* (S). Larger SP graphs are obtained from smaller ones by one of two compositions: a *series composition*, which identifies the South Pole of an SP graph with the North Pole of the other, or a *parallel composition*, which identifies the two North Poles together, and the two South Poles together. For the network flow application we have in mind, we think of SP graphs as directed, with orientation assigned to the edges to allow the flow to move from the North Pole to the South Pole. Figure 1 shows two directed SP graphs, and two directed SP graphs that can be obtained from them by a series and a parallel composition.

2.1. Growth via edge serialization and parallelization

A *series extension* of an edge uv (directed from u to v) is an operation performed by creating a new vertex, say x, removing the directed edge uv and replacing it with the two new directed edges ux and xv. A *parallel doubling* of an edge is performed by just adding another edge with the same endpoints and directed in the same sense. Exercise 11.7 in [9] shows that any SP graph can be obtained by a sequence of serialization and parallelization operations starting at K_2. As an example, the graph on the top right of Figure 1 can be

[1] An abstract definition of a flavor of the SP graph qualifies a graph to be in the family, if it does not contain K_4 as a minor.

https://doi.org/10.1017/S0269964818000505 Published online by Cambridge University Press

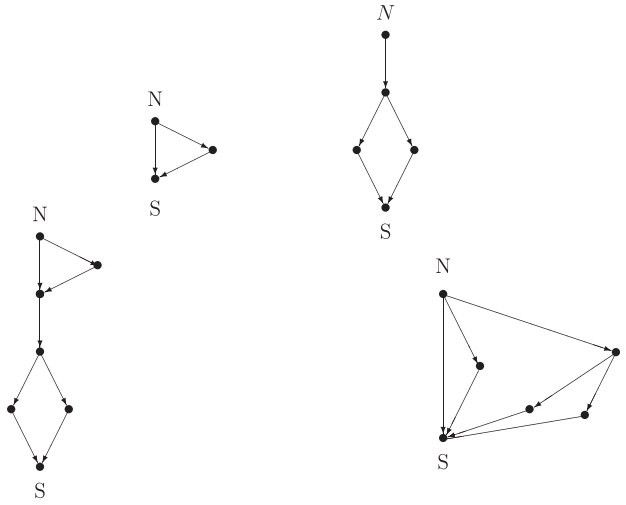


FIGURE **1.** Two directed series–parallel graphs on top, and a graph obtained by a series composition (bottom left), and another graph obtained by their parallel composition (bottom right).

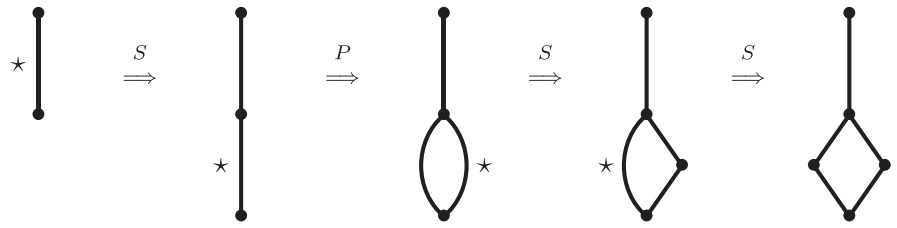


FIGURE **2.** The evolution of a graph by edge serialization (S) and parallelization(P). The chosen edges are starred.

obtained by a sequence of evolutionary steps shown in Figure 2. In Figure 2, the edges selected for serialization/parallelization are indicated by asterisks.

2.2. A spectrum of evolutionary SP graphs

We consider a rather general growth model. Assume we have a *building sequence* $k_1, k_2, k_3, \ldots$ of nonnegative integers, such that

$$k_n \le \tau_0 + \sum_{j=1}^{n-1} k_j, \tag{1}$$

with arbitrary integer $\tau_0 \ge 1$. We start with an arbitrary *seed* SP graph with at least τ_0 edges. Often, this starting seed is K_2, where $\tau_0 = 1$, but occasionally we allow the seed to be an SP graph different from K_2 to conform with a given building sequence, such as one having $k_1 > 1$, in which case the seed must have at least k_1 edges. We grow an SP graph in steps. At step n, we choose k_n edges from the existing graph at random (all subsets of k_n edges being equally likely). Each edge in the selected subset is subjected independently to either a series extension with probability p, or a parallel doubling with probability $q := 1 - p$.

https://doi.org/10.1017/S0269964818000505 Published online by Cambridge University Press

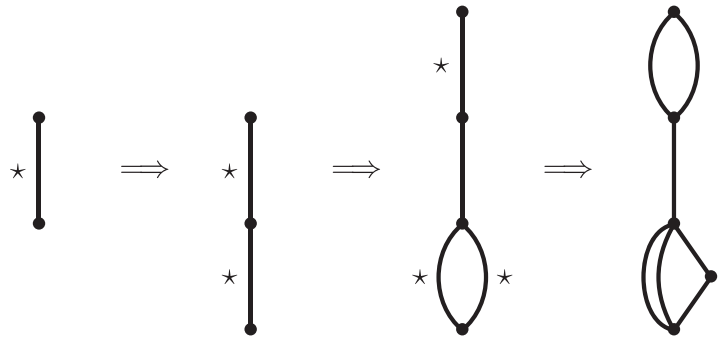


FIGURE **3.** The evolution of a graph by edge serialization (S) and parallelization(P) under the building sequence $k_n = n$. The chosen edges are starred.

The rationale behind the constraint (1) is to make the selection of edges feasible at each step of the growth process: At no point in time does one demand to select a number of edges more than those that exist in the graph. Figure 3 illustrates the growth of an SP starting from K_2, under the building sequence $k_n = n$. The edges selected at one stage are starred.

Henceforth "random" will always mean the model we have just introduced, with multiple edge selection and probabilities p and $q = 1 - p$ for the independent serialization and parallelization of the selected edges.

Let τ_n be the *size* of the graph (in terms of its number of edges) after n steps of evolution. According to the construction, we have

$$\tau_n = \tau_0 + \sum_{j=1}^{n} k_j.$$

In this article, we look at some properties of a random SP graph evolving under a feasible building sequence. Specifically, we look at the degrees of certain types of nodes in the random SP graph: We shall examine the degree of a pole and the number of critical nodes (of outdegree 1).

3. PROBABILITY NOTATION

We use the notation $\mathrm{Ber}(s)$ for a Bernoulli random variable that assumes the value 1 (success) with probability s; otherwise it assumes the value 0 (failure), with probability $1 - s$. We employ the notation $\mathrm{Bin}(j, s)$ for a binomially distributed random variable arising from counting the number of successes among j independent identically distributed Bernoulli experiments, each with success probability s. The notation $\mathrm{Hypergeo}(m, w, s)$ is for a hypergeometric random variable counting the number of white balls in a sample of size s balls taken out of an urn containing w white balls and $m - w$ blue balls.

Technically speaking, such random variables should be defined on the probability space on which the graphs grow. To meet this requirement, we take $\mathrm{Ber}(s)$ to be the indicator $\mathbb{I}_{\{U<s\}}$ of the event that a standard uniform (0,1) random variable, defined on the space of graphs and independent of the growth process, is less than s. Likewise, $\mathrm{Bin}(j, s)$ is the sum of j independent such indicators (and each indicator arises from an independent uniform

https://doi.org/10.1017/S0269964818000505 Published online by Cambridge University Press

(0,1) random variable). We can formally represent the binomial random variable $\mathrm{Bin}(j, s)$ as

$$\mathrm{Bin}(j, s) = \mathbb{I}_{\{U_1<s\}} + \cdots + \mathbb{I}_{\{U_j<s\}},$$

where U_j are independent standard uniform (0,1) random variables defined on the same probability space as the SP graph. Each uniform random variable is independent of all others generated for the evolution, as well as the history of the process.

4. THE DEGREE OF THE NORTH POLE

The number of edges emanating out of the North Pole is a measure of the bottleneckedness at the pole, or the volume of trading and the amount of goods that can be shipped out of the source. (The internal structure of the graph may add further restrictions.) This number is the outdegree of the North Pole (it is also its degree, as this particular node is the only one in the SP graph that does not have any predecessor).

A color code will help phrasing some ideas succinctly. Suppose we colored the edges emanating out of the North Pole with white, and all the other edges with blue. Let W_n be the number of white edges in the graph after n steps of edge additions to the starting seed graph. So, W_n is the outdegree of the North Pole. At the nth stage of evolution, we pick k_n edges from the graph at random. For each selected edge, we generate an independent $\mathrm{Ber}(p)$ random variable. If the Bernoulli random variable associated with an edge realizes the value 1 (failure to increase the degree of its northern end vertex), we serially extend the chosen edge, otherwise it realizes the value 0 (success to increase the degree of its northern end vertex), and we double the edge by parallelization. If the edge is white, and the associated Bernoulli random variable is 1, we are extending an edge connected to the North Pole; this adds one blue directed edge to the graph. If the edge is white, and the Bernoulli random variable is 0, we only double the edge, adding another white edge to the graph. If, instead, we sample a blue edge, we add a blue edge, because no matter what operation is performed on the added edge, it does not change the outdegree of the North Pole.

4.1. The average north polar degree

The number of white edges in the nth sample, X_n, is conditionally the hypergeometric random variable $\mathrm{Hypergeo}\,(\tau_{n-1}, W_{n-1}, k_n)$.

Of X_n a binomially distributed number will be parallelized (converted into double white edges). Specifically, the number of white edges at age n is what it was at age $n-1$ plus an independent $\mathrm{Bin}(X_n, q)$, giving rise to the stochastic recurrence

$$W_n = W_{n-1} + \mathrm{Bin}(X_n, q). \tag{2}$$

We can now develop a recurrence for the average value of the degree of the North Pole. We take the conditional expectation of (2), and obtain

$$\mathbb{E}\big[W_n \,|\, W_{n-1}\big] = W_{n-1} + q\,\mathbb{E}[X_n \,|\, W_{n-1}].$$

Pluggining in the average of the hypgeometric distribution (see [11]) simplifies the recurrence to

$$\mathbb{E}\big[W_n \,|\, W_{n-1}\big] = W_{n-1} + q\frac{W_{n-1}}{\tau_{n-1}}k_n. \tag{3}$$

https://doi.org/10.1017/S0269964818000505 Published online by Cambridge University Press

Taking a double expectation, we obtain a recurrence, which is namely

$$\mathbb{E}[W_n] = \left(1 + \frac{qk_n}{\tau_{n-1}}\right) \mathbb{E}[W_{n-1}].$$

This recurrence is readily available in an iterable form. Unwinding it we obtain an explicit formula.

THEOREM 1: *Suppose $\{k_n\}_{n=1}^{\infty}$ is a building sequence of a family of SP graphs, starting from a SP seed with polar degree W_0. Let W_n be the outdegree (also the degree) of the North Pole of the network after n steps of evolution according to the growth sequence. We then have*

$$\mathbb{E}[W_n] = \prod_{j=1}^{n} \left(1 + \frac{qk_j}{\tau_{j-1}}\right) W_0.$$

To appreciate the wide range of applicability of Theorem 1, we consider a few examples.

EXAMPLE 1 (constant sample size): *Consider the case of constant addition of edges, $k_n = k$, where k is a positive integer. The seed SP graph must have at least k edges in it to enable the first step of evolution. So, $k \le \tau_0$ is required. Of these τ_0 edges in the seed, there are $W_0 \ge 1$ connected to the North Pole.*

In this case, we have $\tau_n = kn + \tau_0$. The average in Theorem 1 is further simplified to

$$\mathbb{E}[W_n] = \left(\prod_{j=1}^{n} \left(1 + \frac{qk}{k(j-1) + \tau_0}\right)\right) W_0 = \frac{\Gamma(\tau_0/k)\,\Gamma(n + q + (\tau_0/k))}{\Gamma(q + (\tau_0/k))\,\Gamma(n + (\tau_0/k))} W_0.$$

The average formula is now amenable to straightforward asymptotic analysis. An asymptotic formula ensues from Stirling's approximation of the ratio of Γ functions:

$$\mathbb{E}[W_n] \sim \frac{\Gamma(\tau_0/k) W_0}{\Gamma(q + (\tau_0/k))}\, n^q, \quad \text{as } n \to \infty.$$

In the case $k = 1$, we retrieve the result in [14]. In this case, an exact distribution of the polar degree is derived in [14].

EXAMPLE 2 (hierarchical lattice): *Suppose the seed is K_2 and at each step, all the edges in the SP graph are independently subjected to evolution. Namely, we have $k_n = 2^{n-1}$. This SP growth produces the so-called hierarchical lattice network [7]. In this case, we have $\tau_n = 2^n$, and $k_n = \tau_{n-1}$. The average in Theorem 1 is then readily simplified to*

$$\mathbb{E}[W_n] = (q+1)^n.$$

EXAMPLE 3 (linear growth): *Examples 1 and 2 deal with the extremes: At one end of the spectrum (Example 1), we have slow constant growth. At the other end of the spectrum (Example 2), we have the fastest possible growth (exponential). Let us consider something in between, a case of linear growth with $k_n = n$, starting from the seed K_2. For simplicity*

https://doi.org/10.1017/S0269964818000505 Published online by Cambridge University Press

of calculation, we take $p = 1/2$. In this case, we have $\tau_n = n(n+1)/2 + 1$. The average in Theorem 1 then becomes

$$\mathbb{E}[W_n] = \prod_{j=1}^{n} \left(\frac{j^2+2}{j^2-j+2} \right) W_0.$$

One has a better chance at analysis, if one deals with factorized quadratic formulas. We have

$$\begin{aligned}
\mathbb{E}[W_n] &= \prod_{j=1}^{n} \frac{(j - i\sqrt{2}\,)(j + i\sqrt{2}\,)}{(j - \frac{1}{2} - \frac{1}{2}i\sqrt{7}\,)(j - \frac{1}{2} + \frac{1}{2}i\sqrt{7}\,)} W_0 \\
&= \frac{\Gamma(n+1-i\sqrt{2}\,)\,\Gamma(n+1+i\sqrt{2}\,)\,\Gamma(\frac{1}{2} - \frac{1}{2}i\sqrt{7}\,)\,\Gamma(\frac{1}{2} + \frac{1}{2}i\sqrt{7}\,)}{\Gamma(n+\frac{1}{2} - \frac{1}{2}i\sqrt{7}\,)\,\Gamma(n+\frac{1}{2} + \frac{1}{2}i\sqrt{7}\,)\,\Gamma(1-i\sqrt{2}\,)\,\Gamma(1+i\sqrt{2}\,)} W_0 \\
&\sim \frac{\Gamma(\frac{1}{2} - \frac{1}{2}i\sqrt{7}\,)\,\Gamma(\frac{1}{2} + \frac{1}{2}i\sqrt{7}\,) W_0}{\Gamma(1-i\sqrt{2}\,)\,\Gamma(1+i\sqrt{2}\,)}\, n;
\end{aligned}$$

in the last line we applied Stirling's approximation for the ratio of Γ functions. The constant can be simplified using the formulas

$$\Gamma(1-iy)\Gamma(1+iy) = \frac{\pi y}{\sinh(\pi y)}, \quad \Gamma\left(\frac{1}{2} - iy\right)\Gamma\left(\frac{1}{2} + iy\right) = \frac{\pi}{\cosh(\pi y)}.$$

Hence, in the linear growth case we asymptotically have

$$\mathbb{E}[W_n] \sim \frac{\sinh\big(\pi\sqrt{2}\,\big)}{\sqrt{2}\,\cosh\big((\pi/2)\sqrt{7}\,\big)}\, W_0\, n = 0.9417541725\ldots W_0 n, \qquad \textit{as} \quad n \to \infty.$$

4.2. Beyond average

The relation (3) invites a martingale to appear in the structure. Writing that equation in the form

$$\mathbb{E}\big[W_n \,|\, W_{n-1}\big] = \left(1 + \frac{qk_n}{\tau_{n-1}}\right) W_{n-1} := g_n W_{n-1},$$

we immediately realize that $W_n^* = W_n/(g_1 g_2 \ldots g_n)$ is a martingale with respect to the natural filtration generated by the evolving SP graphs.

We appeal to the martingale convergence theorem (see Theorem 3.1 in [6]). The theorem states that if X_n is a martingale sequence, such that $\sup_{n\geq 1} \mathbb{E}|X_n| < \infty$, then X_n converges almost surely to a limiting random variable X^* (and in L_1, too). The theorem is not constructive; it does not specify the limit X^*.

Specializing to random SP graphs, we shall see that the North Pole degree, upon appropriate scaling, always converges almost surely. According to Theorem 1, the martingale W_n^* has mean W_0, for all $n \geq 0$. Regardless of the building sequence, we have

$$\frac{W_n}{\prod_{i=1}^{n}(1 + ((qk_i)/(\tau_{i-1})))} \xrightarrow{a.s.} W^*,$$

where W^* is a limiting random variable.

https://doi.org/10.1017/S0269964818000505 Published online by Cambridge University Press

EXAMPLE 4 (constant sample size, revisited): *In Example* 1*, we developed the asymptotics of the scale:* $g_1 \dots g_n \sim cn^q$*, where* c *is* $\Gamma(\tau_0/k)/\Gamma((\tau_0/k)+q)$*. So, we now have two convergence relations*

$$\frac{W_n}{g_1 \dots g_n} \xrightarrow{a.s.} W^*, \quad \textit{and} \quad \frac{g_1 \dots g_n}{n^q} \to c$$

Combining these relations, we see that

$$\frac{W_n}{n^q} \xrightarrow{a.s.} W^{**},$$

for some *limiting random variable* W^{**}*.*

EXAMPLE 6 (hierarchical lattice, revisited): *In the take-all strategy of Example* 2*, the scaling is amenable to asymptotic characterization. In this case, the scale is*

$$g_1 \dots g_n = \prod_{i=1}^{n} \left(1 + \frac{q\tau_{n-1}}{\tau_{n-1}}\right) = (1+q)^n.$$

It follows that

$$\frac{W_n}{(1+q)^n} \xrightarrow{a.s.} W^*,$$

for some *limiting random variable* W^**.*

EXAMPLE 5 (linear growth, revisited): *In Example* 3*, we developed the asymptotics of the scale:* $g_1 \dots g_n \sim 0.9417541725 \dots W_0 n$*. So, we have*

$$\frac{W_n}{n} \xrightarrow{a.s.} W^{**},$$

for some *limiting random variable* W^{**}*.*

5. CRITICAL NODES AND EDGES

Nodes of outdegree 1 are critical in the network. A failure in a sole critical edge (maybe a road) coming out of such a node causes a great disruption in the overall communication across the network; all the paths going through this node are paralyzed. Therefore, an edge out of a node with outdegree 1 is "critical" in some sense.

To assess the seriousness of critical edges, we count their number. We go as usual through the selection of edges, k_n of them at step n. The selected edges are subjected to serialization or parallelization. With probability p, a critical edge in the sample is extended into a path of length 2, both edges in the path are critical; a net gain of one critical edge. Alternatively, with probability q, the chosen critical edge is parallelized into two noncritical edges, and one critical edge is lost. On the other hand, the evolution of a noncritical edge adds a critical edge, when the evolving edge is serialized.

Let C_n be the number of critical edges in the network after n steps of evolution, and A_n be the number of critical edges appearing in the nth sample; thus there are $k_n - A_n$ noncritical edges in the sample. We arbitrarily number the critical edges in the sample with $1, 2, \dots, A_n$ and the noncritical edges with $1, 2, \dots, k_n - A_n$. The evolution of the jth critical edge, $j \in \{1, 2, \dots, A_n\}$, in the sample adds $2B_j - 1$ critical edges, where B_j is a Bernoulli

https://doi.org/10.1017/S0269964818000505 Published online by Cambridge University Press

(p) random variable, that is independent of all else. The evolution of the ℓth noncritical edge, $\ell \in \{1, 2, \ldots, k_n - A_n\}$, in the sample adds B_ℓ^* critical edges, where B_ℓ^* is a Bernoulli (p) random variable, that is independent of all else.

We can now write a stochastic recurrence for the critical edges:

$$C_n = C_{n-1} + \sum_{j=1}^{A_n} (2B_j - 1) + \sum_{\ell=1}^{k_n - A_n} B_\ell^*.$$

Let $\mathcal{F}_n$ be the sigma field generated by the first n steps of multi-edge evolution. Taking conditional expectations, we get

$$\mathbb{E}[C_n \,|\, \mathcal{F}_{n-1}, A_n] = C_{n-1} + \mathbb{E}\left[\sum_{j=1}^{A_n} (2B_j - 1) \,\middle|\, \mathcal{F}_{n-1}, A_n\right] + \mathbb{E}\left[\sum_{\ell=1}^{k_n - A_n} B_\ell^* \,\middle|\, \mathcal{F}_{n-1}, A_n\right].$$

For the expectations involving sums, we need a Wald-like equation (see Theorem 3.3.2 in [16]), for example:

$$\mathbb{E}\left[\sum_{j=1}^{A_n} (2B_j - 1) \,\middle|\, \mathcal{F}_{n-1}, A_n\right] = \mathbb{E}\big[(2B_1 - 1) \,|\, \mathcal{F}_{n-1}, A_n\big] + \cdots + \mathbb{E}\big[(2B_{A_n} - 1) \,|\, \mathcal{F}_{n-1}, A_n\big].$$

The random variables $B_1, \ldots, B_{A_n}$ have the same distribution, giving

$$\mathbb{E}\left[\sum_{j=1}^{A_n} (2B_j - 1) \,\middle|\, \mathcal{F}_{n-1}, A_n\right] = A_n \,\mathbb{E}\big[(2B_1 - 1) \,|\, \mathcal{F}_{n-1}, A_n\big].$$

Furthermore, the variable B_1 is independent of A_n and of the sigma field $\mathcal{F}_{n-1}$, leading us to

$$\mathbb{E}\left[\sum_{j=1}^{A_n} (2B_j - 1) \,\middle|\, \mathcal{F}_{n-1}, A_n\right] = (2p - 1)A_n.$$

Likewise, we have

$$\mathbb{E}\left[\sum_{\ell=1}^{k_n - A_n} B_\ell^* \,\middle|\, \mathcal{F}_{n-1}, A_n\right] = p(k_n - A_n).$$

Putting the parts together and taking double expectation

$$\mathbb{E}[C_n] = \mathbb{E}[C_{n-1}] + (p - 1)\,\mathbb{E}[A_n] + pk_n. \tag{4}$$

5.1. Regularity condition

Under a simple regularity condition on the building sequence, we shall see that the average proportion of critical edges approaches an identifiable limit. The condition is

$$\frac{k_n}{\tau_n} \to \beta \in [0, 1], \quad \text{as } n \to \infty. \tag{5}$$

The condition is quite natural and is met by most sequences that arise in practice. Such a practical sequence would have a systematic generating mechanism. For example, the

https://doi.org/10.1017/S0269964818000505 Published online by Cambridge University Press

constant, linear, and take-all sequences that appear in Examples 1–3 follow this regularity condition, with $\beta = 0$ for the former two building sequences and with $\beta = 1/2$ for the latter.

However, there are conceivably erratic sequences not following the regularity condition. Take, for example, a building sequence like

$$k_n = \begin{cases} 2^{\frac{n-1}{2}}, & \text{if } n \text{ is odd;} \\ 0, & \text{if } n \text{ is even.} \end{cases}$$

For this sequence, $k_{2n}/\tau_{2n} \to 0$, while $k_{2n-1}/\tau_{2n-1} \to 1/2$. The subsequences converge to different limits and the regularity condition does not hold.

5.2. Critical edges: average

The formulation in (4) allows a computation of the exact and asymptotic average of the number of critical edges.

THEOREM 2: *Suppose $\{k_n\}_{n=1}^{\infty}$ is a building sequence of a family of SP graphs, starting from a seed with C_0 critical edges. Let C_n be the number of critical edges in the network after n steps of evolution according to the building sequence. We have*

$$\mathbb{E}[C_n] = p\sum_{i=1}^{n} k_i \prod_{j=i+1}^{n}\left(1 - \frac{qk_j}{\tau_{j-1}}\right) + C_0 \prod_{j=1}^{n}\left(1 - \frac{qk_j}{\tau_{j-1}}\right),$$

and under the condition $k_n/\tau_n \to b \in [0,1]$ (as $n \to \infty$), we have the average limiting proportion

$$\mathbb{E}\left[\frac{C_n}{\tau_n}\right] \to \frac{p}{2-p}.$$

PROOF: In (4), A_n is conditionally a Hypergeo $(\tau_{n-1}, C_{n-1}, k_n)$ random variable, with conditional mean $k_n C_{n-1}/\tau_{n-1}$. So, we can write the recurrence in (4) as

$$\mathbb{E}[C_n] = \left(1 - \frac{qk_n}{\tau_{n-1}}\right)\mathbb{E}[C_{n-1}] + pk_n. \tag{6}$$

This is a linear recurrence of form

$$x_n = h_n x_{n-1} + pk_n,$$

and has a standard solution, yielding the exact expectation as in the theorem statement.

The strategy for the asymptotic part of the statements is twofold: We prove the existence of a limit (under any building sequence) for the proportion of critical edges from the exact solution. We then find the value of the limit from the recurrence under a mild regularity condition.

First, express the expected proportion as

$$\mathbb{E}\left[\frac{C_n}{\tau_n}\right] = p\sum_{i=1}^{n} \frac{k_i}{\tau_n} \prod_{j=i+1}^{n}\left(1 - \frac{qk_j}{\tau_{j-1}}\right) + \frac{C_0}{\tau_n}\prod_{j=1}^{n}\left(1 - \frac{qk_j}{\tau_{j-1}}\right).$$

Let

$$a_n = \sum_{i=1}^{n} \frac{k_i}{\tau_n} \prod_{j=i+1}^{n}\left(1 - \frac{qk_j}{\tau_{j-1}}\right).$$

We show that a_n converges to a limit. For this, it suffices to argue that the sequence a_n is increasing and bounded. Since $k_j \le \tau_{j-1}$, and q is a probability in $(0,1)$, the terms in the

https://doi.org/10.1017/S0269964818000505 Published online by Cambridge University Press

product are positive; the summand adds positive terms, and a_n is an increasing sequence. We further have $a_n < \mathbb{E}[C_n/\tau_n] \leq 1$. Thus, a_n has a limit. Likewise, the product in the part

$$b_n := \frac{C_0}{\tau_n} \prod_{j=1}^{n} \left(1 - \frac{qk_j}{\tau_{j-1}}\right)$$

is decreasing with n, and bounded from below by 0, and thus b_n has a limit. Consequently, $\mathbb{E}[C_n/\tau_n]$ has a limit.

As we established the existence of a limit, we can get its value by taking limits of a recurrence equation for the proportion. From (6), we have the equivalent recurrence

$$\mathbb{E}\left[\frac{C_n}{\tau_n}\right] = \left(1 - \frac{qk_n}{\tau_{n-1}}\right) \mathbb{E}\left[\frac{C_{n-1}}{\tau_{n-1}}\right] \frac{\tau_{n-1}}{\tau_n} + p\frac{k_n}{\tau_n}.$$

While a limit average proportion always exists, it is most easily identified with the aid of the regularity condition (5). Taking limits (the existence of which has been established), we can identify the limit under the regularity condition (5), when $\beta > 0$:

$$c = \left(1 - \frac{q\beta}{1-\beta}\right) c(1-\beta) + p\beta,$$

which we can solve algebraically for c to obtain

$$\lim_{n\to\infty} \mathbb{E}\left[\frac{C_n}{\tau_n}\right] = c = \frac{p}{2-p}.$$

The same limit holds for $\beta = 0$ via a continuity argument: The solution $c = c(\beta)$ may depend on β. If we assume $\beta > 0$ to be very small, we get the equation $c(\beta) = ((p\beta)/(\beta(2-p)))$, taking limits, as $\beta \to 0$, we find $c(0^+) = (p/((2-p)))$. ■

It is curious to note that a limit for $\mathbb{E}[C_n/\tau_n]$ always exists (as in the proof), no matter how bizarre the building sequence is. The regularity condition (5) is only a device to facilitate finding the limit. The same limiting value may appear for an irregular sequence as the one alternating with parity mentioned in Subsection 5.1. In this case, we have

$$\begin{aligned}\mathbb{E}\left[\frac{C_{2n+1}}{\tau_{2n+1}}\right] &= \left(1 - \frac{qk_{2n+1}}{\tau_{2n}}\right) \mathbb{E}\left[\frac{C_{2n}}{\tau_{2n}}\right] \frac{\tau_{2n}}{\tau_{2n+1}} + p\,\frac{k_{2n+1}}{\tau_{2n+1}} \\ &= \left(1 - q\,\frac{2^n}{2^n-1}\right) \mathbb{E}\left[\frac{C_{2n}}{\tau_{2n}}\right] \frac{2^n-1}{2^{n+1}-1} + p\,\frac{2^n}{2^{n+1}-1}.\end{aligned}$$

Taking limits, as $n \to \infty$, we get $\beta = (1-q)\beta + (p/2)$, with solution $\beta = (p/(2-p))$.

The exact solution in Theorem 2 can be unwieldy in its full generality. However, it is possible in some cases, when the building sequence has a systematic generating mechanism in it, to develop an explicit closed-form average and sharper asymptotics. Let us reconsider, for instance, two of the three examples in Subsection 4.1 from the point of view of critical nodes.

EXAMPLE 4 (constant sample size): *Consider $k_n = k$, a constant, where k is a positive integer, and the starting seed has $\tau_0 \geq k$ edges, of which C_0 are critical. In this case,*

https://doi.org/10.1017/S0269964818000505 Published online by Cambridge University Press

$\tau_n = kn + \tau_0$. *The exact expectation in Theorem 2 takes the form*

$$\begin{aligned}\mathbb{E}[C_n] &= p\sum_{i=1}^{n} k \prod_{j=i+1}^{n}\left(1-\frac{qk}{k(j-1)+\tau_0}\right) + C_0\prod_{j=1}^{n}\left(1-\frac{qk}{k(j-1)+\tau_0}\right),\\ &= pk\sum_{i=1}^{n}\prod_{j=i+1}^{n}\frac{j-1+(\tau_0/k)-q}{j-1+(\tau_0/k)} + C_0\prod_{j=1}^{n}\frac{j-1+(\tau_0/k)-q}{j-1+(\tau_0/k)}\\ &= pk\sum_{i=1}^{n}\frac{\Gamma(i+(\tau_0/k))\,\Gamma(n+(\tau_0/k)-q)}{\Gamma(n+(\tau_0/k))\,\Gamma(i+(\tau_0/k)-q)} + C_0\frac{\Gamma((\tau_0/k))\,\Gamma(n+(\tau_0/k)-q)}{\Gamma((\tau_0/k)-q)\,\Gamma(n+(\tau_0/k))}.\end{aligned}$$

Using a known formula for the sum, we obtain

$$\begin{aligned}\mathbb{E}[C_n] &= \frac{pk}{2-p}\left(n+\frac{\tau_0}{k}-\frac{\Gamma((\tau_0/k)+1)\,\Gamma(n+(\tau_0/k)-q)}{\Gamma(n+(\tau_0/k))\,\Gamma((\tau_0/k)-q)}\right)\\ &\quad + C_0\frac{\Gamma((\tau_0/k))\,\Gamma(n+(\tau_0/k)-q)}{\Gamma(n+(\tau_0/k))\,\Gamma((\tau_0/k)-q)}\\ &= \frac{pk}{2-p}\,n + \frac{p\tau_0}{2-p} + O\left(\frac{1}{n^q}\right).\end{aligned}$$

EXAMPLE 5 (hierarchical lattice): *The seed for a hierarchical lattice is* K_2, *and* $\tau_0 = 1$, $C_0 = 1$. *When we develop all the edges, we have* $k_n = \tau_{n-1} = 2^{n-1}$. *The exact expectation takes the simple form*

$$\begin{aligned}\mathbb{E}[C_n] &= p\sum_{i=1}^{n} 2^{i-1}\prod_{j=i+1}^{n}(1-q) + \prod_{j=1}^{n}(1-q),\\ &= p\sum_{i=1}^{n} 2^{i-1}p^{n-i} + p^n\\ &= \frac{p}{2-p}\,2^n - \frac{p^{n+1}}{2-p} + p^n\\ &\sim \frac{p}{2-p}\,2^n, \quad \text{as } n\to\infty,\end{aligned}$$

as in [5].

References

1. Barabási, A. (2016). *Network science*. Cambridge, UK: Cambridge University Press.
2. Bernasconi, N., Panagiotou, K., & Steger, A. (2008). On the degree sequences of random outerplanar and series-parallel graphs. *Lecture Notes in Computer Science* 5171: 303–316.
3. Drmota, M. (2008). *Random trees: an interplay between combinatorics and probability*. Springer.
4. Drmota, M., Giménez, O., & Noy, M. (2010). Vertices of given degree in series-parallel graphs. *Random Structures & Algorithms* 36: 273–314.
5. Feng, Y., Mahmoud, H., & Rüschendorf, L. (2016). Degree profile of hierarchical lattice networks. *Probability in the Engineering and Informational Sciences* 31: 60–82.
6. Hall, P. & Heyde, C. (1980). *Martingale limit theory and its application*. New York: Academic Press Inc.
7. Hambly, B. & Jordan, J. (2004). A random hierarchical lattice: the series-parallel graph and its properties. *Advances in Applied Probability* 36: 824–838.

https://doi.org/10.1017/S0269964818000505 Published online by Cambridge University Press

8. Hibbard, T. (1962). Some combinatorial properties of certain trees with applications to searching and sorting. *Journal of the ACM* 9: 13–28.
9. Hofri, M. & Mahmoud, H. (2016). *Algorithmics of nonuniformity: tools and paradigms*. Boca Raton, FL: Chapman-Hall.
10. Hofri, M., Li, C., & Mahmoud, H. (2016). On the combinatorics of binary series-parallel graphs. *Probability in the Engineering and Informational Sciences* 30: 224–260.
11. Kendall, M., Stuart, A., & Ord, K. (1987). *Advanced theory of statistics*. Vol. I. Distribution Theory. Oxford University Press.
12. Knuth, D. (1973). *The art of computer programming*. Vol. 3. Reading, MA: Addison-Wesley.
13. Mahmoud, H. (1992). *Evolution of random search trees*. New York: John Wiley & Sons.
14. Mahmoud, H. (2013). Some node degree properties of series-parallel graphs evolving under a stochastic growth model. *Probability in the Engineering and Informational Sciences* 27: 297–307.
15. Mahmoud, H. (2014). Some properties of binary series-parallel graphs. *Probability in the Engineering and Informational Sciences* 28: 565–572.
16. Ross, S. (1996). *Stochastic processes*. New York: Wiley.
17. Szpankowski, W. (2011). *Average case analysis of algorithms on sequences*. New York: Wiley.
18. Van der Hofstad, R. (2017). *Random graphs and complex networks*. Vol. 1. Cambridge, UK: Cambridge University Press.

https://doi.org/10.1017/S0269964818000505 Published online by Cambridge University Press